Dell PowerEdge R750 servers featuring Dell PowerEdge RAID Controllers (PERC 11) delivered stronger Apache Hadoop big data performance while maintaining high availability

The Dell solution outperformed an HPE ProLiant DL380 Gen9 server with an HPE Smart Array P440ar Controller

Overview

Big data applications have become central to business operations across a wide range of industries. To contend with the large volume of data they now collect, store, and access, companies are taking a close look at different storage options. As they do, they often consider two requirements: performance and data availability. While it was once necessary to choose between the high performance of non-volatile memory express (NVMe®) storage and the data redundancy of RAID, with the Broadcom-based Dell™ PowerEdge™ RAID Controller 11 (PERC 11), this is no longer the case: You can have both.

We conducted benchmark testing of an Apache™ Hadoop® big data workload on the current-generation Dell PowerEdge R750 server with Dell PowerEdge RAID Controller 11 (PERC 11). To quantify the advantage of such a solution over a previous-gen compute-and-storage solution, we chose the HPE ProLiant DL380 Gen9 server with HPE Smart Array P440ar. Both solutions used RAID for redundancy, but the storage interfaces differed. The Dell used NVMe solid-state drives (SSDs), while the HPE solution had serial attached SCSI (SAS) SSDs. We used a disk-intensive TeraSort workload from the HiBench suite of benchmarks to measure the performance of each solution. While both solutions provided the data protection that RAID offers, the Dell solution completed the Apache Hadoop workload in 27 percent less time than the HPE solution. It also achieved 36 percent greater throughput than the HPE solution.

Storage options for big data: Balancing strong performance with high availability

Companies increasingly turn to big data workloads to solve business problems such as understanding customer habits and behavior, maintaining electronic health records, and detecting fraud. In a 2022 survey of executives, 97.0 percent had invested in big data initiatives, and 73.7 percent said their organizations had appointed a Chief Data Officer (CDO), up from 12 percent in 2012.[1]

By definition, big data is voluminous. Companies must collect, manage, and store great quantities of information and—to provide a good user experience and/or get the most value from real-time insights—workloads must be able to access and process that data quickly. With so much data at play, storage becomes an essential consideration for companies as they select hardware platforms to run their vital workloads. Two primary requirements for storage are fast performance and availability. Companies seek storage that can quickly put actionable insights into the hands of decision-makers. At the same time, it is a fact of life that storage media occasionally fails, and no company wants to risk the potentially very large expense of losing vital business data.

For several decades, to maximize availability, companies have employed RAID solutions. RAID stands for redundant array of independent disks. As the name suggests, RAID solutions, which can be either software- or hardware-based, let systems manage storage disks in such a way that if one disk fails, no data loss occurs.

Table 1 presents the seven RAID levels in widest use.

Table 1: Comparison of seven popular RAID levels. Source: Principled Technologies.

RAID levels	RAID 0	RAID 1	RAID 5	RAID 6	RAID 10	RAID 50	RAID 60
Description	Striping only	Mirroring	Striping with parity	Striping with double parity	Mirroring and striping	Striping and distributed parity	Striping and double parity
Minimum disks required	2	2	3	4	4	6	8
Relative read performance	High	High	High	High	High	High	High
Relative write performance	Sequential: Very high Random: Very high	Sequential: Medium Random: Medium	Sequential: High Random: Low	Sequential: High Random: Low	Sequential: Medium Random: Medium	Sequential: High Random: Medium	Sequential: High Random: Low-medium
Relative cost difference (usable GB/$)	Lower	Higher	Medium	Medium	Higher	Medium	Medium
Data redundancy	✗	✓	✓	✓	✓	✓	✓
Max drive loss	0	1	1	2	1+	1+	2+
Common use cases	Temporary files, application caches	OS, database logs	Data warehousing or reporting, video streaming, file servers	Data warehousing or reporting, video streaming, file servers	Database data files, application servers	Larger databases, file servers	Larger databases, file servers

Until more recently, hardware-based RAID controllers have been compatible only with older, slower storage protocols such as SAS and SATA. Thus, customers have had to sometimes prioritize either performance or availability. Meanwhile, enterprise storage technology has evolved for years, from serial advanced technology attachment (SATA) storage using spinning disks to SAS storage using SSDs to NVMe storage using SSDs and then to SSDs using the even faster NVMe protocol (see Table 2). Each new technology introduces a new level of speed at a new set of price points. The older, slower storage approaches do not go away, but remain cost-effective ways for organizations to store types of data that they don't need to access rapidly. For example, a hospital must maintain medical records of former patients, but can tolerate a delay in retrieving these. This makes this type of data a good a candidate for a more archival approach using slower, more affordable storage.

Companies seeking the best performance and high availability can now take advantage of new RAID controllers, such as the Dell PowerEdge RAID controllers (PERC 11) that are compatible with the NVMe protocol. These new devices offer companies a best-of-both-worlds option, where buyers need not choose between performance and data availability.

Table 2: Comparison of three storage interfaces. Note: We base this table on a similar TechTarget table.[2]

Interface	Input/output operations per second (IOPS)	Throughput	Latency	Queues	Commands per queue
SATA	60,000 to 100,000	6 Gbps	Below 1 millisecond (ms) to over 100 ms	1	32
SAS	200,000 to 400,000	12 Gbps	Below 100 microseconds (μs) to over 100 ms	1	256
NVMe	200,000 to 10,000,000	32 Gbps (Gen3x4) 64 Gbps (Gen4x4)	Below 10 μs to 225 μs	65,535	65,536

How the Dell PERC11, using Broadcom technologies, lets you choose both performance and availability

The emergence of NVMe storage several years ago brought significant storage performance gains, but no technology, no matter how fast, can remove the need for redundancy options. Many core business applications, backed by legacy relational databases or NoSQL databases, require redundancy at the storage layer as a best practice. But many businesses also could benefit from the faster storage throughput and I/O capabilities of NVMe, so this created an issue. Early in the NVMe transition, those who wanted to use RAID NVMe disks on servers had limited options to ensure that their data would be reliably available; the main options were software RAID or software-defined-storage solutions. This is because historically, RAID controllers assumed the presence of spinning disks, with caches on the controller to compensate for slower mechanically spinning technologies. Also, because NVMe disks connected directly to the PCIe bus, storage controller and server manufacturers had to rethink their controller implementations, a process that naturally took some time to make its way into products.

The Dell PERC11, using the Broadcom RAID-on-chip (ROC) SAS3916 chipset, solves this issue. PERC11 remains compatible with SAS and SATA technologies and handles NVMe disks, as well. According to Dell, PERC 11 RAID controllers offer support for PCIe Gen 4, support for hot-swapping devices, non-volatile cache, secure enterprise key manager security, and more.[3]

The Dell-Broadcom partnership on the RAID controller brings flexibility and performance together in the storage subsystem, allowing for system and application architects to use RAID 0, 1, 5, 6, 10, 50, and 60. This gives database architects the power to make new decisions on file placement and redundancy without sacrificing performance.

About the Broadcom RAID-on-chip (ROC), SAS3916 chipset

The PERC processor in the Dell PERC11 is a Broadcom ROC SAS3916 chipset. Broadcom based this chip on its Fusion-MPT architecture and says the chip "delivers enhanced performance and power reductions over previous generations. The ROC features Tri-Mode SerDes technology that enables a seamless operation of SAS, SATA or NVMe storage devices from any system design.

"The 16-port Tri-Mode ROC device provides SAS data transfer rates of 12, 6 and 3Gb/s per lane and SATA data transfer rates of 6 and 3 Gb/s per lane. The high-port count ROC helps eliminate storage bottlenecks with eight PCI Express® lanes and complies with the PCIe 4.0 specification, offering up to 3 million IOPS (JBOD mode) and up to 2.4 million IOPS in RAID (random reads)."[4]

About the Dell PowerEdge R750 server

The Dell PowerEdge R750 is a full-featured, general-purpose 2U rack server featuring 3rd Gen Intel® Xeon® Scalable processors. According to Dell, the PowerEdge R750 is purpose-built to optimize application performance and acceleration with PCIe Gen 4 compatibility, eight channels of memory per CPU, and up to 24 NVMe drives.[5] It also includes "I/O bandwidth and storage to address data requirements – ideal for: traditional corporate IT, database and analytics, virtual desktop infrastructure, AI/ML, and HPC."[6]

To learn more about the Dell PowerEdge R750, check out the spec sheet at https://i.dell.com/sites/csdocuments/Product_Docs/en/poweredge-R750-spec-sheet.pdf.

How the explosion of data has driven the development of distributed databases

Traditional relational databases have been around for nearly half a century and have improved in countless ways. However, the underlying paradigm of how relational systems model data has remained largely consistent, with machines often needing to "scale-up," or become faster via hardware upgrades, to improve performance. In the last few decades, with the tidal waves of data brought on by internet, mobile, IoT, and other technologies, new clustered and distributed "scale-out" database systems have emerged with the goal of processing expansive amounts of data, both structured (as in relational systems) and unstructured (such as documents, pictures, text, and so on).

Hadoop is one such distributed system, comprising the MapReduce engine, the Hadoop Distributed File System (HDFS), Name Nodes, and Data Nodes. Clusters can be quite large in production, but the key to HDFS is its ability to break apart a very large data problem for processing.

According to the Apache wiki, organizations using Apache Hadoop include eBay, Facebook, Hulu, Spotify, Twitter, and dozens of smaller companies and educational institutions. Applications range from reporting/analytics and machine learning to search optimization to matching dating profiles to content generation and data aggregation.[7]

Putting the Dell PowerEdge R750 server featuring Dell PowerEdge RAID Controllers (PERC 11) to the test

To prove the benefits of using the latest Dell server-and-storage controller technology, we chose to deploy two small, virtualized Hadoop environments: one on the Dell PowerEdge R750 server and one on the HPE ProLiant DL380 Gen9 server. The PowerEdge R750 server had a Dell PERC 11 RAID controller, while the HPE ProLiant DL380 Gen9 server had an HPE Smart Array P440ar Controller. Both servers used Linux® VMs and SSDs, with the Dell system also using NVMe SSDs. We chose the fastest drives each controller supported: NVMe for the Dell PERC 11 RAID controller and SAS SSDs for the HPE Smart Array P440ar Controller.

Table 3 provides aspects of the server configurations we tested. For more detailed configuration information, see the science behind the report.

Table 3: System configurations we used in performance testing. Source: Principled Technologies.

Server configuration information	Dell PowerEdge R750	HPE ProLiant DL380 Gen9
Hardware		
Processors	2x Intel Xeon Gold 6348 28 cores each, 2.6 GHz	2x Intel Xeon E5-2650 12 cores each, 2.2 GHz
Storage controller	PERC H755N Front, 8GB cache	Smart Array P440ar, 2GB cache
Disks	6x 1.6TB Dell Ent NVMe v2	6x 960GB SAS Toshiba PX05SVB096Y
Total memory in system (GB)	256	192
Operation system name and version/ build number	VMware® ESXi®, 7.0.3, 20036589	VMware ESXi, 7.0.3, 19482537
Software		
VM operating system	CentOS	
Benchmarking tools		
Hadoop big data performance	TeraSort benchmark, part of the HiBench suite of benchmarks	

We set up and configured the Dell PowerEdge R750 server remotely in a Dell lab; we installed and configured the HPE ProLiant DL380 Gen9 in the PT lab. We installed and configured the latest HPE-customized VMware vSphere® 7.3 on the DL380 server. We installed the hypervisor on two SSD drives and attached six SSD drives to the respective RAID controller on each server. We created a RAID 5 logical drive on these six SSD drives. We deployed a Hadoop cluster with one manager node and four workers on each server. We configured the RAID 5 logical drive as storage for Hadoop Distributed Filesystem.

Below, we present an overview of the steps we carried out. A detailed step-by-step methodology appears in the science behind the report.

1. Rack and cable servers, verify BIOS and firmware levels on each server, and install vCenter Server.
2. Configure a 2-disk RAID 1 volume and a 6-disk RAID 5 volume on the servers.
3. Install VMware vSphere, and format datastores on each server.
4. Create VMs running CentOS on each server.
5. Install Apache Hadoop and Spark on the CentOS VMs, and create manager and worker nodes.
6. Load and configure HiBench suite with TeraSort dataset, and run tests.

Apache Hadoop performance

To measure big data performance, we employed the TeraSort workload from the HiBench suite. In this workload, the TeraGen function generates input data, the TeraSort function uses MapReduce for sorting, and the TeraValidate function validates the output of the sorted data.[8] We selected this tool because it provides insight into the performance of Hadoop clusters and stresses the storage subsystem. The goal of testing was to generate performance data showing both run time and throughput on each platform. We ran the TeraSort workload three times and report the median of three runs. During testing, our experts relied on other performance data to confirm that the two platforms were functioning as we expected them to and that the configurations were comparable.

Figure 1 shows how long each solution took to complete the TeraSort workload. The Dell PowerEdge R750 server with Dell PERC 11 RAID controller took 4 minutes and 13 seconds, which is 27 percent less time than the 5 minutes and 47 seconds necessary on the HPE ProLiant DL380 Gen9 server with the HPE Smart Array P440ar Controller.

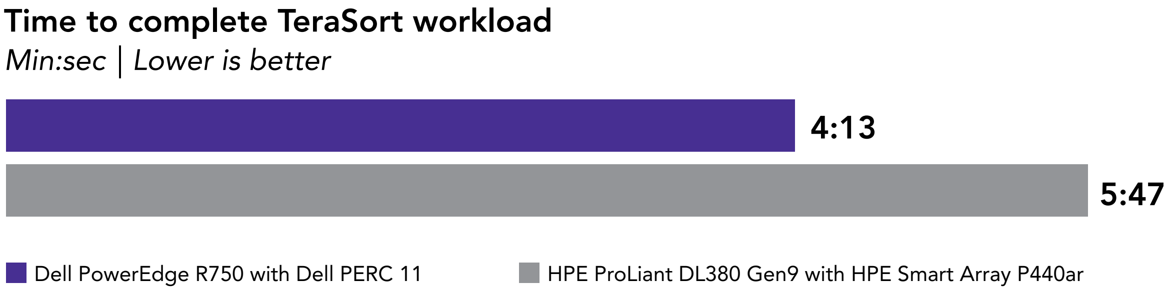

Figure 1: Time required for both solutions to complete a TeraSort workload on Apache Hadoop. Lower is better. Source: Principled Technologies.

Figure 2 compares the throughput rate each solution achieved while completing the TeraSort workload. The Dell PowerEdge R750 server featuring the Dell PERC 11 RAID controller delivered 36 percent more gigabytes per second (GB/s) than the HPE solution with the HPE Smart Array P440ar Controller. A higher throughput rate reflects that the storage solution can process more data in a fixed amount of time.

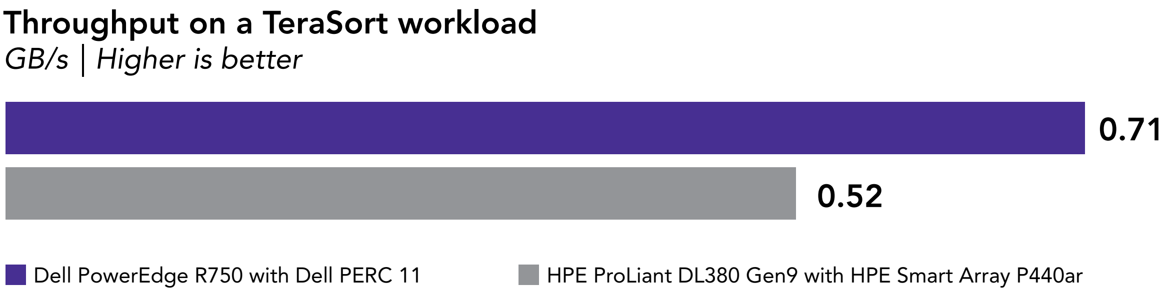

Figure 2: Throughput for both solutions during a TeraSort workload. Higher is better. Source: Principled Technologies.

Why was performance on the Dell solution higher?

With any system upgrade as we have in this scenario, performance advantages are, of course, likely due to multiple factors. The advantages of the newer Dell server in this test are clear, and based on performance data, we can attribute the strong advantage of the Dell PowerEdge R750 with Dell PowerEdge RAID Controller 11 to its use of NVMe storage, the updated storage controller with greater amounts of cache, and newer processors.

Conclusion

Organizations of all sizes have incorporated big data applications into their workflows, and rely on them daily. The enormous volume of information that companies now contend with drives the need for effective storage solutions. These solutions must support strong performance by delivering speedy access to data, which helps companies make critical business decisions in a timely manner. In addition, effective storage solutions protect data and keep it available even if individual storage components stop working.

We ran a disk-intensive TeraSort big data workload on two server-and-storage solutions. Both solutions used RAID for redundancy, but only one of them used high-speed NVMe storage media. The current-generation Dell PowerEdge R750 server with a Dell PERC 11 RAID controller and NVMe storage outperformed the previous-generation HPE ProLiant DL380 Gen9 server with an HPE Smart Array P440ar Controller. The Dell solution completed a disk-intensive TeraSort workload in 27 percent less time and achieved a 36 percent greater throughput rate. These results show that by selecting the Dell PowerEdge R750 server with a Dell PERC 11 RAID controller, companies no longer need to choose between the data protection that comes with true redundant hardware RAID solutions and the performance benefits of the fastest NVMe drives. The Dell-Broadcom solution lets companies have both.

1. New Vantage Partners, "Data and AI Leadership Executive Survey 2022, Executive Summary of Findings," accessed December 8, 2022, https://c6abb8db-514c-4f5b-b5a1-fc710f1e464e.filesusr.com/ugd/e5361a_2f859f3457f24cff9b-2f8a2bf54f82b7.pdf.
2. TechTarget, "NVMe," accessed April 28, 2023, https://www.techtarget.com/searchstorage/definition/NVMe-non-volatile-memory-express.
3. Dell, "Dell PowerEdge RAID Controller 11 User's Guide PERC H755 adapter, H755 front SAS, H755N front NVMe, H755 MX adapter, H750 adapter SAS, H355 adapter SAS, H355 front SAS, H350 adapter SAS, H350 Mini Mo," accessed November 14, 2022, https://www.dell.com/support/manuals/en-pa/perc-h755/perc11_ug/technical-specifications-of-perc-11-cards?guid=guid-aaaf8b59-903f-49c1-8832-f3997d125edf&lang=en-pa.
4. Broadcom, "SAS3916 12Gb/s SAS Tri-Mode RAID-on-Chip (ROC)," accessed May 1, 2023, https://www.broadcom.com/products/storage/raid-on-chip/sas-3916.
5. Dell, "Dell EMC PowerEdge R750 Spec Sheet," accessed November 14, 2022, https://i.dell.com/sites/csdocuments/Product_Docs/en/poweredge-R750-spec-sheet.pdf.
6. Dell, "Dell EMC PowerEdge R750 Spec Sheet."
7. Apache, "Powered by Apache Hadoop," accessed May 2, 2023, https://cwiki.apache.org/confluence/display/HADOOP2/PoweredBy.
8. Ahmed, N., Barczak, A.L.C., Susnjak, T. et al. "A comprehensive performance analysis of Apache Hadoop and Apache Spark for large scale data sets using HiBench," Journal of Big Data, accessed April 28, 2023, https://doi.org/10.1186/s40537-020-00388-5.

Read the science behind this report at https://facts.pt/3m5epzN ▶

Principled Technologies®

Facts matter.®

This project was commissioned by Dell Technologies.